Molecular characterization of the Hansenula polymorpha FLD1 gene encoding formaldehyde dehydrogenase

Baerends, RJS; Sulter, GJ; Jeffries, TW; Cregg, JM; Veenhuis, M; Jeffries, Thomas W.; Cregg, James M.

Published in:
Yeast

DOI:
10.1002/yea.805

Yeast
Yeast 2002; **19**: 37–42.
DOI: 10.1002/yea.805

Yeast Sequencing Report

Molecular characterization of the *Hansenula polymorpha FLD1* gene encoding formaldehyde dehydrogenase

Richard J. S. Baerends[1],†, Grietje J. Sulter[1], Thomas W. Jeffries[2], James M. Cregg[3] and Marten Veenhuis[1]*

[1] *Eukaryotic Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands*

[2] *Institute for Microbial and Biochemical Technology, Forest Products Laboratory, US Department of Agriculture, One Gifford Pinchot Drive, Madison, WI 53705-2398, USA*

[3] *Keck Graduate Institute of Applied Life Sciences, 535 Watson Drive, Claremont, CA 91711, USA*

*Correspondence to: M. Veenhuis, Eukaryotic Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, PO Box 14, 9750 AA Haren, The Netherlands. E-mail: m.veenhuis@biol.rug.nl

†Current address: Molecular Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands.

Received: 24 May 2001
Accepted: 19 August 2001

Abstract

Glutathione-dependent formaldehyde dehydrogenase (FLD) is a key enzyme required for the catabolism of methanol as a carbon source and certain primary amines, such as methylamine as nitrogen sources in methylotrophic yeasts. Here we describe the molecular characterization of the *FLD1* gene from the yeast *Hansenula polymorpha*. Unlike the recently described *Pichia pastoris* homologue, the *H. polymorpha* gene does not contain an intron. The predicted *FLD1* product (Fld1p) is a protein of 380 amino acids (ca. 41 kDa) with 82% identity to *P. pastoris* Fld1p, 76% identity to the FLD protein sequence from *n*-alkane-assimilating yeast *Candida maltosa* and 63–64% identity to dehydrogenase class III enzymes from humans and other higher eukaryotes. The expression of *FLD1* is strictly regulated and can be controlled at two expression levels by manipulation of the growth conditions. The gene is strongly induced under methylotrophic growth conditions; moderate expression is obtained under conditions in which a primary amine, e.g. methylamine, is used as nitrogen source. These properties render the *FLD1* promoter of high interest for heterologous gene expression. The availability of the *H. polymorpha FLD1* promoter provides an attractive alternative for expression of foreign genes besides the commonly used alcohol oxidase promoter. The sequence has been deposited in the GenBank/NCBI data library under Accession No. AF364077.

Keywords: heterologous gene expression; methylotrophic yeast; methanol metabolism

Introduction

Heterologous production of biologically active proteins has become an important tool in fundamental and applied research. During the last decade methylotrophic yeasts have gained increasing interest, both for fundamental (van der Klei and Veenhuis, 1996; Subramani, 1998) and applied purposes (Cereghino and Cregg, 2000; Gellisen, 2000). Currently, the *H. polymorpha* system provides a remarkably versatile technology platform for heterologous gene expression (Gellisen and Veenhuis, 2001). The *Hansenula* toolbox includes a variety of host strains, various sites available for integration and several strong promoters (reviewed by van Dijk *et al.*, 2000). These promoters originate from inducible genes encoding enzymes/proteins involved in: (a) methanol metabolism, e.g. alcohol oxidase (AO), dihydroxyacetone synthase (DHAS) and formate dehydrogenase (FMD) (Faber *et al.*, 1995; Gellisen *et al.*, 1991; Hollenberg *et al.*, 1997; Janowicz *et al.*, 1988); (b) oxidation of primary amines (Zwart *et al.*, 1983); (c) nitrate assimilation (Avila *et al.*, 1998); or (d) an acid phosphatase (Pongdara *et al.*, 1998). Also, strong constitutive promoters have been identified in *H. polymorpha*, e.g. the genes encoding a plasma membrane-bound ATPase (Cox *et al.*, 2000) or the translation elongation factor-1α (Baerends *et al.*, 1997; Kiel and Veenhuis, unpublished results).

Recently, the cloning and characterization of the *P. pastoris FLD1* gene and its promoter (P_{FLD1}) was reported (Shen *et al*., 1998). The gene encodes formaldehyde dehydrogenase (Fld1p), a crucial enzyme in the catabolism of methanol (Veenhuis *et al*., 1983). During methanol metabolism, peroxisomal alcohol oxidase mediates the first oxidation step of methanol to generate formaldehyde and, as a side product, hydrogen peroxide, which is decomposed by peroxisomal catalase to oxygen and water. A portion of the formaldehyde generated from methanol enters the cell biosynthetic pathway to form cell constituents (van Dijken *et al*., 1976; Douma *et al*., 1985), while another portion leaves the peroxisome and is oxidized via formate (by Fld1p) to carbon dioxide (by FMD). These reactions generate reducing power in the form of NADH, thereby serving as the primary source of energy during growth on methanol (Veenhuis *et al*., 1983). Furthermore, Fld1p is thought to play an important role in the protection of the cell from toxic formaldehyde that accumulates when excess methanol is present in the medium (Sibirny *et al*., 1990). Fld1p is also involved in the metabolism of formaldehyde generated from other sources, such as primary amines (e.g. methylamine and choline) used as nitrogen source for growth (Zwart *et al*., 1980).

The *H. polymorpha FLD1* gene was isolated by functional complementation of the *P. pastoris fld1* mutant (Shen *et al*., 1998). Here, we report the molecular characterization of this gene. The availability of the gene allows us to use its promoter for controlled production of other genes of interest, besides the currently available promoters.

Experimental procedures

Strain, media and complementation analysis

The *Pichia pastoris fld1 his4* strain MS105 was cultivated as described previously (Shen *et al*., 1998). Complementation analysis was performed as detailed in Cregg and Russell (1998).

DNA procedures

Escherichia coli DH5α (*supE*44 Δ*lacU*169 (ϕ80*lacZ*ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1) (Sambrook *et al*., 1989) was used for plasmid construction and amplification. Cells were grown on LB medium supplemented with the appropriate antibiotics. For growth on agar plates, the media were supplemented with 1.5% granulated agar. Standard recombinant DNA manipulations (Sambrook *et al*., 1989) were carried out as detailed previously. Endonuclease restriction enzymes and biochemicals were obtained from Roche Diagnostics Nederland BV (Almere, The Netherlands) and used as detailed by the manufacturer. Sequencing of DNA was performed by BaseClear (Leiden, The Netherlands), using universal or specific primers in order to obtain the complete nucleotide sequence of the *FLD1* gene and its upstream promoter region (by primer walking).

Subcloning of *H. polymorpha FLD1*

The plasmid pYG2 containing *H. polymorpha* genomic DNA, which was previously described to complement the *P. pastoris fld1* mutant (Shen *et al*., 1998). In order to locate the *FLD1* gene on the insert, pYG2 was subjected to detailed restriction analysis and four subclones were analysed for their ability to complement the *fld1* mutation. The smallest complementing subclone, a 2.4 kb *SphI* fragment of pYG2, was cloned into *SphI*-digested pUC19 (New England Biolabs, Leusden, The Netherlands). The resulting vector, pHOR1-2, was used for initial sequencing (single-stranded) in order to determine the exact position of the gene on the genomic DNA fragment. Since pHOR1-2 contained only a small *FLD1* promoter fragment, a 3.5 kb genomic DNA fragment from pYG2 was subcloned as an *Asp718* fragment into *Asp718*-digested phagemid pBluescript II KS^+ (pBSII KS^+; Stratagene Inc., San Diego, CA), resulting in pHOR6-1. Subsequently, a 3070 bp genomic fragment of pHOR6-1 was sequenced double-stranded.

Computer analysis

For DNA and amino acid sequence analysis the PC-GENE™ program release 6.70 (IntelliGenetics, Mountain View, CA) was used. Genbank database searches were performed using the BLAST algorithm (Altschul *et al*., 1997). A multiple amino acid sequence alignment was done using the ClustalX program (Thompson *et al*., 1997).

Results and discussion

H. polymorpha formaldehyde dehydrogenase (Fld1p) is a glutathione-dependent enzyme, essential for

the metabolism of formaldehyde generated from methanol or some unusual nitrogen sources (e.g. primary amines, choline) used for growth (Veenhuis *et al.*, 1983; Zwart *et al.*, 1980). Deletion of the gene results in impairment of the cells to grow on methanol as sole carbon source (Zwart *et al.*, 1980). We used this property to clone the *H. polymorpha FLD1* gene by functional complementation of a *P. pastoris fld1* mutant, using an *H. polymorpha* genomic bank. After transformation, methanol$^+$ mutants were selected for further analysis. The complementing activity, indicated by the restoration of the ability of the *P. pastoris fld1* mutant to grow on methanol, was found to reside within a 7.2 kb genomic insert of pYG2 (Shen *et al.*, 1998). To identify the *H. polymorpha FLD1* gene within this sequence, the insert was subjected to restriction analysis and various subclones were made. These subclones were analysed for their ability to complement the *P. pastoris fld1*mutant (Figure 1). The smallest complementing fragment, a 2.4 kb *Sph*I fragment of pYG2, was used for initial sequencing. This fragment contains two putative open reading frames (ORFs). One of these, ORF1, showed the strongest similarity to *Saccharomyces cerevisiae* Snq2p (54%), the other displayed a very high identity to *P. pastoris* formaldehyde dehydrogenase (Fld1p, see below).

Since the 2.4 kb *Sph*I fragment contains only a small *FLD1* promoter fragment, a larger fragment of pYG2 was cloned and sequenced (Figure 1). The DNA sequence of this fragment displayed two ORFs. Besides the *FLD1* gene, an additional ORF was found which has a strong similarity to a *Schizosaccharomyces pombe* ORF encoding a probable peptidyl prolyl *cis-trans*-isomerase (p*PPI*; Accession No. T39728). Unlike the *P. pastoris FLD1* gene, its homologue in *H. polymorpha* did not contain an intron (Shen *et al.*, 1998). The *FLD1* ORF is 1143 bp long and is predicted to encode a protein of 380 amino acids with a calculated molecular mass of 40 655 Da.

The Genbank/NCBI database was searched for other proteins with amino acid sequence similarity to *H. polymorpha* Fld1p. The sequence of the putative *H. polymorpha* Fld1p showed the highest identity (82%), as expected for that of glutathione-dependent Fld1p from *P. pastoris* (Figure 2). Furthermore, the protein displayed a strong similarity (76%) with Fld from the yeast *Candida maltosa*. This yeast assimilates *n*-alkanes and the Fld enzyme is thought to be involved in the

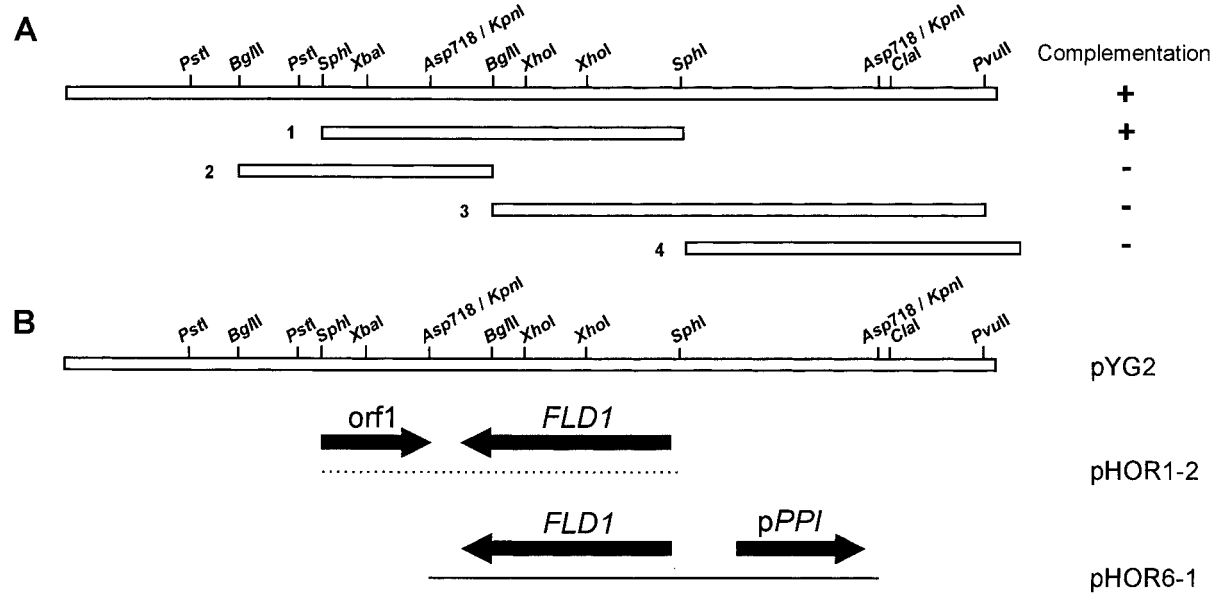

Figure 1. Restriction enzyme map and subcloning of the 7.2 kb genomic insert of pYG2. Various subclones were analysed for their ability to complement the *P. pastoris fld* mutant (A). Sequencing of the complementing subclone 1 (in pHOR1-2) revealed that the fragment contains two open reading frames (see text) (B). The complementation analysis demonstrates that the *fld1*-complementing activity is linked to the *FLD1* gene and not to ORF1 (in pHOR1-2) or p*PPI* (pHOR6-1)

Figure 2. Comparison of the deduced amino acid sequences of *H. polymorpha* (HpFld1p), *Pichia pastoris* (PpFld1p) and *Candida maltosa* (CmFdh1p) formaldehyde dehydrogenase proteins with alcohol dehydrogenase class III enzymes of human and fruitfly (HsAdh5p resp. DmAdhIII)

protection from the toxic effects of formaldehyde (Sasnauskas *et al*., 1992). The *H. polymorpha* Fld1p also showed 63–64% identity with dehydrogenase class III enzymes from humans and other higher eukaryotes (Holmquist and Vallee, 1991; Koivusalo *et al*., 1989; Giri *et al*., 1989).

In *P. pastoris*, the *FLD1* promoter was demonstrated to present an attractive alternative to the commonly-used alcohol oxidase promoter. *P. pastoris* P_{FLD1} was shown to be a highly regulatable promoter that is capable of producing a heterologous protein at levels equal to or higher than those produced by P_{AOX1} (Shen *et al*., 1998). Detailed biochemical analysis have demonstrated that the *H. polymorpha FLD1* gene is also strictly regulated (Eggeling and Sahm, 1978; Zwart *et al*., 1980). When cells are grown on glucose/ammonium sulphate, no formaldehyde dehydrogenase activity is measured, while a shift of cells to media containing methanol as carbon source or primary amines such as methylamine results in strong induction of the gene and high enzyme activities (Eggeling and Sahm, 1978; Zwart *et al*., 1980).

So far, we have made use of the endogenous alcohol oxidase promoter for overexpression of endogenous or heterologous genes in *H. polymorpha* (van Dijk *et al*., 2000). In order to obtain a stable and high expression strain the constructed overexpression cassette is usually integrated into the genomic P_{AOX} locus at a high copy number (Baerends *et al*., 1997; Faber *et al*., 1995). Recently, we found that subsequent overexpression of two different constructs at the same P_{AOX} locus resulted in diminished production levels when compared to strains containing a similar copy number of only one of the two constructs (Otzen and Kiel, unpublished results). It is therefore advantageous to have an additional strong inducible promoter available. Construction of a suitable vector containing the P_{FLD1} promoter, in order to facilitate overexpression of genes of interest, is now under way.

Acknowledgements

RJSB was supported by Grant BIO4-08-0543 from the European Community.

References

Altschul SF, Madden TL, Schaffer AA, *et al.* 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Avila J, Gonzalez C, Brito N, Siverio JM. 1998. Clustering of the YNA1 gene encoding a Zn(II)2Cys6 transcriptional factor in the yeast *Hansenula polymorpha* with the nitrate assimilation genes YNT1, YNI1 and YNR1, and its involvement in their transcriptional activation. *Biochem J* **335**: 647–652.

Baerends RJS, Salomons FA, Faber KN, Kiel JAKW, van der Klei IJ, Veenhuis M. 1997. Deviant Pex3p levels affect normal peroxisome formation in *Hansenula polymorpha*: high steady-state levels of the protein fully abolish matrix protein import. *Yeast* **13**: 1437–1448.

Cereghino JL, Cregg JM. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev* **24**: 45–66.

Cox H, Mead D, Sudbery P, Mark Eland R, Mannazzu I, Evans L. 2000. Constitutive expression of recombinant proteins in the methylotrophic yeast *Hansenula polymorpha* using the *PMA1* promoter. *Yeast* **16**: 1191–1203.

Cregg JM, Russell KA. 1998. Transformation. In *Methods in Molecular Biology: Pichia Protocols,* Higgins DR, Cregg JM (eds). Humana: Totowa, NJ; 27–39.

Douma AC, Veenhuis M, de Koning W, Evers ME, Harder W. 1985. Dihydroxyacetone synthase is localized in the peroxisomal matrix of methanol-grown *Hansenula polymorpha*. *Arch Microbiol* **143**: 237–243.

Eggeling L, Sahm H. 1978. Derepression and partial insensitivity to carbon catabolite repression of the methanol dissimilating enzymes in *Hansenula polymorpha*. *Eur J Appl Microbiol* **5**: 197–202.

Faber KN, Harder W, AB G, Veenhuis M. 1995. Review: methylotrophic yeasts as factories for production of foreign proteins. *Yeast* **11**: 1331–1344.

Gellissen G. 2000. Heterologous protein production in methylotrophic yeasts. *Appl Microbiol Biotechnol* **54**: 741–750.

Gellissen G, Janowicz ZA, Merckelbach A, *et al.* 1991. Heterologous gene expression in *Hansenula polymorpha*: efficient secretion of glucoamylase. *Biotechnology (NY)* **9**: 291–295.

Gellissen G, Veenhuis M. 2001. The methylotrophic yeast *Hansenula polymorpha*: its use in fundamental research and as a cell factory. *Yeast* **18**: i-iii.

Giri R, Krug JF, Kozak C, *et al.* 1989. Cloning and comparative mapping of human class III alcohol dehydrogenase cDNA. *Biochem Biophys Res Commun* **164**: 453–460.

Higgins DR, Cregg JM. 1998. *Methods in Molecular Biology: Pichia Protocols.* Humana: Totowa, NJ; 103.

Hollenberg CP, Gellissen G. 1997. Production of recombinant proteins by methylotrophic yeasts. *Curr Opin Biotechnol* **8**: 554–560.

Holmquist B, Vallee BL. 1991. Human liver class III alcohol and glutathione dependent formaldehyde dehydrogenase are the same enzyme. *Biochem Biophys Res Commun* **178**: 1371–1377.

Janowicz ZA, Merckelbach A, Eckart M, Weydemann U, Roggenkamp R, Hollenberg CP. 1988. Expression system based on the methylotrophic yeast *Hansenula polymorpha*. *Yeast* **4S**: 155.

Koivusalo M, Barmann M, Uotila L. 1989. Evidence for the identity of glutathione-dependent formaldehyde dehydrogenase and class III alcohol dehydrogenase. *FEBS Lett* **257**: 105–109.

Pongdara A, Merckelbach A, Keup P, Gellissen G, Hollenberg CP. 1998. Cloning and characterization of the gene encoding a repressible acid phosphatase (PHO1) from the methylotrophic

yeast *Hansenula polymorpha*. *Appl Microbiol Biotechnol* **50**: 77–84.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press: New York.

Sasnauskas K, Jomantiene R, Januska A, Lebediene E, Lebedys J, Janulaitis A. 1992. Cloning and sequecing analysis of a *Candida maltosa* gene which confers resistance to formaldehyde in *Saccharomyces cerevisiae*. *Gene* **122**: 207–211.

Shen S, Sulter G, Jeffries TW, Cregg JM. 1998. A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Gene* **216**: 93–102.

Sibirny AA, Ubiyvovk VM, Gonchar MV, *et al.* 1990. Reactions of direct formaldehyde oxidation to CO_2 are non-essential for energy supply of yeast methylotrophic growth. *Arch Microbiol* **154**: 566–575.

Subramani S. 1998. Components involved in peroxisome import, biogenesis, proliferation, turnover, and movement. *Physiol Rev* **78**: 171–188.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **24**: 4876–4882.

van Dijk R, Faber KN, Kiel JAKW, Veenhuis M, van der Klei I. 2000. The methylotrophic yeast *Hanenula polymorpha*: a versatile cell factory. *Enzyme Microb Technol* **26**: 793–800.

van Dijken JP, Oostra-Demkes GJ, Otto R, Harder W. 1976. S-formylgluthathione: the substrate for formate dehydrogenase in methanol-utilizing yeasts. *Arch Microbiol* **111**: 77–83.

van der Klei IJ, Veenhuis M. 1996. Peroxisome biogenesis in the yeast *Hansenula polymorpha*: a structural and functional analysis. *Ann NY Acad Sci* **804**: 47–59.

Veenhuis M, van Dijken JP, Harder W. 1983. The significance of peroxisomes in the metabolism of one-carbon compounds in yeasts. *Adv Microb Physiol* **24**: 1–82.

Zwart K, Veenhuis M, van Dijken JP, Harder W. 1980. Development of amine oxidase-containing peroxisomes in yeast during growth on glucose in the presence of methylamine as the sole source of nitrogen. *Arch Microbiol* **126**: 117–126.

Zwart K, Veenhuis M, Plat G, Harder W. 1983. Characterization of glyoxysomes in yeasts and their transformation into peroxisomes in response to changes in the environmental conditions. *Arch Microbiol* **136**: 28–38.